

Boston University

OpenBU

<http://open.bu.edu>

Theses & Dissertations

Boston University Theses & Dissertations

2015

Quality of life and depressive symptoms as predictors of participant adherence in a randomized trial conducted among older adults at risk of mobility disability

<https://hdl.handle.net/2144/15619>

Boston University

BOSTON UNIVERSITY
SCHOOL OF MEDICINE

Thesis

**QUALITY OF LIFE AND DEPRESSIVE SYMPTOMS AS PREDICTORS OF
PARTICIPANT ADHERENCE IN A RANDOMIZED TRIAL CONDUCTED
AMONG OLDER ADULTS AT RISK OF MOBILITY DISABILITY**

By

HAO ZHU

B.S., University of Illinois, 2012

Submitted in partial fulfillment of the
Requirements for the degree of
Master of Arts
2015

© 2015 by

HAO ZHU

All rights reserved

Approved By

First Reader

Janice Weinberg, ScD.

Director of M.S. in Clinical Investigation

Second Reader

Thomas G. Travison, PhD.

Senior Scientist and Director of Biostatistics

Institute for Aging Research, Hebrew Seniorlife

Third Reader

Ilean Isel Isaza, PhD.

Study Data Coordinator

Institute for Aging Research, Hebrew Seniorlife

ACKNOWLEDGMENTS

I would like thank Dr. Thomas G. Trivison for his guidance throughout this project. You are my teacher, my friend and my mentor. You gave me knowledge, and more importantly, the opportunity to grow. I would also like to thank my committee members, Dr. Janice Weinberg and Dr. Ilean Isel Isaza for your brilliant comments and suggestions. You made me think and helped me make this thesis a lot better.

Then, I would like to thank Dr. Weinberg and Ms. Stacey Hess Pino in the Master of Arts in Clinical Investigation program at Boston University. Thank you for your understandings when things were not going smoothly as planned.

At the end, I would like to appreciate my wife, Siyi Gu, who were always by my side when I need to work at night. Thank you for your support.

**QUALITY OF LIFE AND DEPRESSIVE SYMPTOMS AS PREDICTORS OF
PARTICIPANT ADHERENCE IN A RANDOMIZED TRIAL CONDUCTED
AMONG OLDER ADULTS AT RISK OF MOBILITY DISABILITY**

HAO ZHU

ABSTRACT

Poor adherence is an issue in clinical trials with a striking magnitude and negative impact. Recent studies indicate that two widely used clinical screening tools, Short Form (SF-36) Health Survey and Center for Epidemiologic Studies Depression Scale (CES-D) can be used as risk stratifiers to identify participants who require extra interaction to stay adherent in specific population. There is, however, little evidence to support the implication of these tools in elderly individuals with mobility limitations. These individuals may be particularly vulnerable to the risks of poor adherence in the context of a demanding interventional scheme. This study fulfills this gap by analyzing the quality of life and adherence data from the VIVE2 study, which is a double-blinded randomized explanatory clinical trial assessing the benefits of nutritional supplements and daily exercises to elderly patients with mobility limitation. The preliminary results showed that in clinical studies targeting elderly population with mobility limitations, the summary scores of SF-36 trended to have weak and nonsignificant association with a decreased risk of poor adherence to both exercise completion and product consumption while CES-D has no association.

TABLE OF CONTENTS

TITLE.....	i
COPYRIGHT PAGE	ii
READER APPROVAL PAGE	iii
ACKNOWLEDGMENTS	iv
ABSTRACT	v
TABLE OF CONTENTS	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS	xi
INTRODUCTION	1
The Importance of Participant Adherence to Clinical Research	1
Factors thought to influence Adherence	2
Rationale for Prediction of Poor Adherence at Trial Baseline.....	5
Rationale for Choice of Quality of Life as potential risk-stratifier on Poor Adherence.....	6
Adherence and Aging.....	8
Background of Mobility Disability in Elderly Population	10
Limited Evidences in Elderly Subjects at Risk of Mobility Disability	10

The VIVE2 Study.....	12
Study Objectives	13
Primary Objectives	13
Secondary Objectives	14
Summary.....	14
METHODS.....	16
Overview of the VIVE2 Study	16
Study Product Description.....	16
Physical Activity Program Description	16
Screening, Randomization and Blinding Protocols.....	17
Inclusion and Exclusion Criteria for the VIVE2 Study.....	18
Measures	18
SF-36 Measures.....	18
CES-D Measures	19
400 M Walk Measures.....	20
Nutritional Intervention Adherence	20
Exercise Program Adherence.....	21
Definition of Risk of Poor Adherence.....	21
Statistical Analysis Plan	22

Primary Analysis.....	22
Secondary Analyses.....	23
RESULTS.....	25
Description of trial sample.....	25
Distribution of Adherence statistics	26
Association of SF-36 and CES-D Scores with Poor Adherence.....	27
Association of Test Scores with Continuous Measures of Nonadherence	28
Associations of SF-36 Subscales with Poor Adherence.....	29
Risk Prediction Model for Poor Adherence.....	30
Other Factors	31
DISCUSSION	33
Interpretation of the Results	33
Study Limitations.....	36
APPENDIX I. Extra Terminologies.....	40
Inclusion Criteria for the VIVE2 Study	40
Exclusion Criteria for the VIVE2 Study.....	40
APPENDIX II. Extra Tables	43
BIBLIOGRAPHY.....	45
CURRICULUM VITAE	50

LIST OF TABLES

TABLE I. Components of SF-36 Tests.....	19
TABLE II. Baseline Overall Descriptive Characteristics (N=150)	25
TABLE III. Multivariate Logistic Regression Models Demonstrating the Association of Baseline SF-36 and CES-D Scores with Poor Adherence among Subjects in the VIVE2 Study.....	28
TABLE IV. Multivariate Linear Regression Models Demonstrating the Association of Baseline SF-36 and CES-D Scores with Risk of Nonadhering Events (%) among Subjects in the VIVE2 Study.....	29
Table V. Associations between each SF-36 Subscale and Poor Adherence	30
Table VI. Multivariate Logistic Regression Model Assessing the Associations of Each SF-36 Subscale and Poor Adherence	31
Table A1. Baseline Descriptive Characteristics among Randomization Groups.....	42
Table A2. Multivariate Logistic Regression Model Assessing Other Factors Affecting Risk of Poor Adherence	43

LIST OF FIGURES

Figure 1. Boxplots showing distribution of exercise nonadherence score by site.

According to the protocol, participants with 40% proportionate nonadherence
(above blue line) are considered noncompliant. 27

LIST OF ABBREVIATIONS

BMI	Body Mass Index
BP	Bodily pain
CES-D	Center for Epidemiologic Studies Depression Scale
CT	Computed Tomography
DXA	Dual Energy X-Ray Absorptiometry
GFR	Glomerular Filtration Rate
GH	General health perceptions
HRQOL	Health-related quality of life
ITT	Intention to treat
MCS	Mental Component Summary
MH	Mental health
MMSE	Mini-mental state examination
MNA-SF	Mini Nutrition Assessment Short Form
NYHA	New York Heart Association
PCS	Physical Component Summary
PF	Physical Functioning
PP	Per-protocol
RCT	Randomized Clinical Trial
RE	Emotional role functioning
RP	Physical role functioning
SAP	Statistical Analysis Plan

SD	Standard Deviation
SF	Social role functioning
SF-36	Short Form Health Survey
SPPB	Short Physical Performance Battery
VT	Vitality
WHO	World Health Organization

INTRODUCTION

Participant adherence, or compliance, is an important determinant of the outcome of both clinical practices and clinical trials. In the context of delivery of medical care, a 2003 World Health Organization (WHO) guidebook gave the definition of adherence as “the extent to which a person’s behavior – taking medication, following a diet, and /or executing lifestyle changes, corresponds with the agreed recommendations from a health care provider”¹. In clinical research, the related concept of adherence describes the extent to which a subject’s behavior corresponds with what is anticipated by the consented study protocol. As discussed below, a high rate of participant adherence to the consented protocol is essential in generating valid research findings. This thesis uses an analysis of a randomized clinical trial (RCT) to investigate the hypothesis that in the context of a complex illness (loss of physical functioning), generalized measures of quality of life and participant well-being will be predictive of adherence to a demanding intervention schedule.

The Importance of Participant Adherence to Clinical Research

In ideal cases, a good clinical study should be built on a trust relationship between investigators and subjects. After subjects are well consented about the procedure, risks and benefits of the study, they are expected to strictly follow the study protocol. In practice, however, cases of poor adherence are widely found in the field of clinical research. In fact, reports have demonstrated that the

adherence rates in clinical trials ranges from 43% to 78% among subjects receiving treatments for different chronic diseases². A few reviews also reported that in developed countries, the adherence rates in patients treated for chronic diseases are around 50%³, while in developing countries, the rates might be even lower due to the poor availability and accessibility of healthcare resources¹. It is clear, therefore, that poor adherence can jeopardize the validity of clinical research in any setting.

In clinical trials, poor adherence may be a critical problem because it may reduce the statistical power of the study, increase its cost, and may have grave implications for the applicability of its conclusions to the background population. When subjects are not compliant, the benefits of intervention might be underdetermined and the risks poorly understood. If the overall adherence rates in a study were very low, it would be reasonable to question whether the conclusion of the study were in any way useful or applicable outside the confines of the study itself. There are also ethical implications, because if a clinical study cannot produce valuable knowledge, the risks taken by subjects were probably wasted for no reason⁴.

Factors thought to influence Adherence

Due to the striking magnitude of poor adherence and its negative impact on clinical research, in the past decades, many studies have investigated the inner mechanism of non-compliant behaviors. Before the 1990s, most of the

studies were focused on the association between adherence and patients' personality traits^{5,6}. The results, however, were not consistent and sometimes contradictory. Since then, it has been recognized that adherence is actually a very complex behavioral process determined by several interacting factors¹. These factors include characteristics of the disease, characteristics of the therapy, attributes of the patient and the patient's environment.

A systematic analysis in 2008 reviewed 102 cases from 2095 related articles and categorized all the potential risk factors for nonadherence into two groups: "hard" factors and "soft" factors⁷. Hard factors tend to be related and difficult to change, either necessary complexities of the intervention or system-level obstacles or difficulties. These include therapy-related factors (administration methods, duration and side effects), healthcare system problems (accessibility and satisfaction), disease characteristics and healthcare expenditures. Generally, subjects tend to avoid following the therapy when the administration methods are complicated, the duration is long, or the side effects are noticeable. Their adherence rates will increase as the accessibility and satisfaction of healthcare system increases or the cost of the therapy decreases. Patients with a more severe disease also tend to have a higher rate of adherence. These hard factors are quantifiable but are relatively static. It is true that problems in healthcare system are amenable to study and improvement with great investment of resources. Increasing the accessibility and satisfaction of

healthcare system, however, requires a long time commitment and a strong social support. As a result, most of these hard factors are not easy to be amenable to management in clinical trials.

The soft factors are the subject-level obstacles that may be realistically (relatively speaking) accommodated or, perhaps, improved in the context of a clinical study. These include subjects' demographic factors, psycho-social factors such as subjects' beliefs, attitude (depression, anxiety, fears or anger about the illness) and their motivation for therapy. Studies have found that females, the individual's increased levels of education, and marital status are positively associated with a greater adherence to intervention protocols⁷. In younger populations, age is inversely related to adherence, whereas in older populations, age tends to be associated with a higher rate of adherence⁷. Regarding subjects' psycho-social factors, it was reported that subjects' beliefs about the causes and meaning of illness, and motivation to follow the therapy were strongly positively related to their adherence rates⁸⁻¹³. It also has been found that in elderly populations, major depression or anxiety was associated with a higher risk of poor adherence¹⁴. As compared with "hard factors," these soft factors may be more meaningful for assessment in a clinical trial as they are more likely to vary from participant to participant and the information about most of these factors are usually collected at the baseline part of the studies. It may be useful, therefore, to

consider these factors as potential risk stratifiers for nonadherence within a particular trial.

Rationale for Prediction of Poor Adherence at Trial Baseline

An important lesson learnt over the past 40 years is that poor adhering subjects need to be supported by investigators to achieve the best results¹. After cases of poor adherence occurred, it would never be a good solution to “blame” participants for nonadherence behavior. Blaming will not help fix the bad data but actually damages the subject-investigator relationship and makes the subjects become less likely to be compliant in the future^{15,16}. One of the most powerful methods found to increase adherence rates is the good interaction and communication between investigators and subjects^{15,17}. Thus, in clinical research, investigators should take the responsibility of keeping subjects from poor adherence by providing adequate education before study starts and maintaining good communications with subjects during the study.

Perhaps the biggest challenge of doing clinical research is that the resources are often severely limited. It is easy to set a goal that investigators should have substantial direct communications with subjects throughout the study. In many cases, however, it is very hard to achieve this goal, especially when the sample population is large. It is difficult for investigator to direct attention to all each individual participant, and needless to say there are some participants who require additional attention to stay compliant. As a result, there

is a need of to better identify sub-groups of participants with a higher risk of poor adherence at the beginning of the study. Then, the study investigators can have an opportunity to better support them throughout the intervention period.

Rationale for Choice of Quality of Life as potential risk-stratifier on Poor Adherence

For a useful baseline measure to predict risks of poor adherence in clinical research, we propose the following requirements:

- The measure should be widely applied and available;
- The measure should broadly incorporate the patient related risk factors for poor adherence described above;
- The measure must be feasibly deployable at baseline of a clinical research study.

Limited research suggests that it is possible to use the Short Form (SF-36) Health Survey¹⁸, which measures health-related quality of life (HRQOL), and the Center for Epidemiologic Studies Depression Scale (CES-D)¹⁹, which measures the presence of depressive symptomatology, as tools to predict poor adherence. A study conducted in 2011 found that the SF-36 measures were significantly associated with decreased adherence among urban African Americans with severe, poorly controlled hypertension²⁰. Another study, which evaluated hazardous alcohol use and depressive symptomatology among HIV-infected patients in Nigeria, reported that an elevated CES-D score is associated with

decreased adherence as well²¹. Below we give a brief overview of these assessments.

The SF-36 test is a well-validated short-form health survey intended to evaluate physical/mental health and general well-being [citation for SF-36 design]. It consists of 36 questions and generates an 8-subscale profile, which includes vitality (VT), physical functioning (PF), bodily pain (BP), general health perceptions (GH), physical role functioning (RP), emotional role functioning (RE), social role functioning (SF) and mental health (MH). Questions in these 8 subscales can be reorganized into two categories: Physical Component Summary (PCS) and Mental Component Summary (MCS). All of these subscales and component summaries can be evaluated independently for different purposes. Due to its high accuracy and flexibility, SF-36 became a popular general clinical instruments for many purposes, since it was first developed in 1992¹⁸. It has been translated into more than 170 languages and is widely used around the world²².

The CES-D scale is a screening assessment for major depression and depressive disorders, which is widely used in clinical research as an indicator of the presence and severity of depressive symptoms. It consists of 20 questions assessing the 9 symptoms of depression, including sadness (dysphoria), loss of interest (anhedonia), irregular appetite, decreased sleep quality, difficulties in thinking and concentration, feelings of guilt or lack of self-worth, excessive

tiredness (fatigue), increased undirected movement (agitation) and suicidal ideation. The CES-D scale is one of the most widely used medical instrument in the field of psychiatric epidemiology, and is commonly employed in aging research and in other disciplines²³.

Both SF-36 and CES-D tests are commonly used as baseline tests in clinical trials, address in broad terms the patient-related factors of poor adherence discussed above, and can be readily administered at the start of a research study. These therefore meet our broad criteria for eligible risk stratifies. As noted above, there currently exist only a few studies assessing the feasibility of using these two tests as predictors of poor adherence and most of these studies are focusing on specific populations, and none of these have addressed physical function as an outcome. There are still large amounts of knowledge about the implication of these methods in many other different populations waiting to be explored.

Adherence and Aging

An unexplored population is the seniors. Even though many studies have found that in elderly populations, age is associated with a slightly higher adherence rates⁷, it is undeniable that certain subgroups of older subjects are more vulnerable to some barriers to adherence¹. Examples of such subgroups include subjects with multiple chronic diseases or multi-morbidity, subjects with

memory difficulties and subjects with mobility limitations. The latter are the target population of the analyses reported here.

Importantly, many elderly subject have multiple chronic diseases. As a result, they are usually recommended to take multiple medications together with the study medication simultaneously. It has been found that this excess burden of medications – or polypharmacy - may confuse subjects, especially when subjects are required to take these medications multiple times a day²⁴. Multiple medications increase the risk of missing medications, dosing errors, and drug interactions, and therefore might conceivably increase the risks of poor adherence in an RCT.

Also, mobility limitation, which is one of the most common health issues among elder subjects, is another important factor to consider for elderly subjects at risk of poor adherence. Mobility limitations prevent elderly subjects from coming to clinical visits or prevent them from coming to visits on time¹. First of all, it increases the chance of getting missing data. At the same time, it decreases the chance for subjects to communicate and build close bonds with investigators, which have been proven to be one of the significant determinants of poor adherence. Thus, elderly subjects in clinical research need particular attention in clinical management to maintain compliant in clinical trials.

Background of Mobility Disability in Elderly Population

Mobility limitation and loss of functional ability in aging are predictive of decreased quality of life, disability and premature mortality. They are also associated with greater resource utilization and increased health care costs^{25–28}. It is therefore critically important to prevent loss of mobility in aging, and numerous interventions including physical activity, resistance training, mind/body exercise, nutritional or behavioral adaptations, pharmacologic agents and other interventions are currently the focus of testing as function-promoting therapies and for the prevention of injury^{29–33}. The project described here is one of such interventions.

Limited Evidences in Elderly Subjects at Risk of Mobility Disability

As discussed above, elderly subjects with mobility limitations are among the subgroups of elderly population at larger risks of poor adherence. Currently, however, there are only very limited number of studies focusing on adherence in this population^{34–38}. A good example of such studies is the Lifestyle Interventions and Independence for Elders Pilot (LIFE-P) study. LIFE-P study is a well-organized, 12-month clinical trial designed to compare the effects of physical activity intervention (PA) and a health education control on risk of disability in elderly population³³. According to this study, adherence to physical activity is associated with a higher Short Physical Performance Battery (SPPB) score³⁸, which is a validated indicator of a better general mobility and physical

functioning³⁹. This finding confirms the important role of physical activity adherence for elderly people to maintain a good mobility.

A secondary analysis of the LIFE-P study in 2007 also tried to address the predictors of adherence to physical activity in this study³⁴. To my knowledge, this is the only study assessing the predictors of adherence in the elderly population with mobility limitations. The results of this analysis indicated that baseline demographic variables, disease burden, self-reported symptoms, physical functioning and social cognitive variables could predict 10% of the variance in the adherence rate during intervention. If the adherence rate during the baseline test is also considered inside this model, the amount of explained variance increased from 10% to 21%. The investigators also report that the associations were not very consistent as different phases of the study, which indicated risks of poor adherence, might change with time.

Thus, currently, regarding the purpose of validate the use of baseline SF-36 and CES-D scores as clinical instruments to predict subject adherence during the study for elderly subject with mobility limitations, the existing scientific evidences are very limited. As a result, we conduct this study using data from the VIVE2 study, a recent clinical trial addressing the mobility issues in elders, to fulfill this need.

The VIVE2 Study

The “Efficacy of Nutritional Supplementation on Physical-activity Mediated Changes in Physical Functioning Older Adults at Risk for Mobility Disability” (or VIVE2) study is a randomized, multi-centered (U.S. and Sweden), double-blind clinical trial that compares the effects of a 6 month (28 weeks) structured physical activity program with a daily nutritional supplement (NESTLÉ® “Senior II”) or placebo on changes in functional limitations in older adults aged ≥ 70 years with mobility limitations. The primary outcome is the difference between the walking speed of a 400 M walk at baseline and that after 6 months. Other outcomes include the functional limitations (3 month 400 M walking speed vs that at baseline, Stair climb, and SPPB score), body composition (Dual Energy X-Ray Absorptiometry (DXA) and Computed Tomography (CT)), peak torque and power measure, nutritional status (Mini Nutrition Assessment Short Form (MNA-SF), quality of life (SF-36) and depressive symptoms (CES-D).

The primary analytic goal for the VIVE2 study is to assess whether the nutritional intervention can out-perform placebo in producing exercise-induced changes in physical function. The sub-study described here utilizes the VIVE2 study database to assess evidence of association between quality of life and adherence to intervention, without regard to randomization. Per the VIVE2 protocol, a participant is classified as “Poor adherent” if (s)he participates in

fewer than 60% of planned interactions per protocol over the 24-week intervention period.

Study Objectives

Primary Objectives

The primary objectives of this study are to assess whether the baseline physical and mental SF-36 summary scores (PCS & MCS) and CES-D scale are associated with the odds of being classified as “Poor adherent” in an elderly population with mobility limitations. Based on past literature^{20,21}, we hypothesized that a higher baseline PCS score of the SF-36 test would be associated with an increased higher odds of being “Poor adherent,” while the MCS score of the SF-36 test probably would have no such association. We also hypothesized that baseline level of depressive symptoms (assessed by CES-D Scale) would be probably positively associated with the odd of being “Poor adherent”.

Because all participants in the VIVE2 study receive a structured exercise program (the same regardless of randomization), and each receives (according to randomization) either a nutritional intervention or placebo, the project described here considers both the exercise and nutritional interventions in independent analyses. As noted above, the threshold of 60% adherence applies to both the exercise and nutritional activities, independently.

Secondary Objectives

Analysis of “proportionate” adherence

Instead of being counted as a binary (“Good adherent”/“Poor adherent”) variable, poor adherence will also be analyzed as a quantitative variable to provide more information about the effects of the baseline tests’ scores on the probability of a non-adhering event to occur. For the purposes of this analysis, the adherence outcome is computed as the simple proportion of planned activities completed by the participant, with exercise and nutritional components again treated individually.

SF-36 Subscale Score Analyses

Similarly, the association of each SF-36 subscale (PF, RP, BP, GH, VT, SF, RE and MH) and poor adherence will be analyzed.

Multiple Regression Model of Poor Adherence

Finally, a risk-prediction model for poor adherence will be developed taking into account other factors including age, sex and other demographic factors and baseline measures.

Summary

Poor adherence is an issue in clinical trials with a striking magnitude and negative impact. Past research indicates the only way to solve this issue is to provide study participants more supports, rather than placing responsibility on

them alone. In principle, the use of standard surveys as potential risk stratifiers of poor adherence may help in anticipating the subgroups of subjects at greatest need of additional attention at little to no additional cost of detection.

Investigators might then use the limited resources available to them to better support participants who require extra interaction to stay compliant. Recent studies indicate that SF-36 and CES-D are reasonable candidates in specific populations, but there is little existing evidence to support the implication of these tools in elderly individuals with mobility limitations, who may be particularly vulnerable to the risks of poor adherence in the context of a demanding interventional scheme. This study will address this gap by analyzing quality of life and adherence data from the VIVE2 study, which is a double-blinded randomized clinical trial assessing the benefits of nutritional supplements and daily exercises to elderly patients with mobility limitation.

METHODS

Overview of the VIVE2 Study

In this study, 150 subjects aged 70 years or older at risk of mobility disability were randomized to receive daily either the nutritional supplement (NESTLÉ® “Senior II”) or a placebo for a period of 6 months. During the study, all participants were also required to take a 3-time per week multi-model physical activity training program (strength, balance and stretching). Clinical measures were taken at the baseline, 3 month visit and 6 months visit.

Study Product Description

The nutritional supplement used in this study, NESTLÉ® “Senior II”, is a 119 mL (4 Oz) nutritional beverage. The formulas consist of 20 g high quality whey protein, 800 IU vitamin D, and 350 mg Ca. It can provide 150 kcal of energy with 6.5 g carbohydrate, and 6 g lipids. The placebo was a sweetened and flavored beverage serving at a size of 119 mL with 20-30 calories. The packages of the nutritional supplement and the placebo were identical.

Physical Activity Program Description

Subjects in both intervention arm and placebo arm were also required to participate in a multi-model physical exercise program specifically designed by Tufts University and Nestlé. The purpose of this activity program was to introduce subjects to the strength, stretching, and balance portions of the program in a safe and effective manner. These center-based exercise instruction sessions were

hosted three times per week and were supervised by nurses and investigators. The exercise duration was approximately one hour which included a warm-up period, 30 minutes of aerobics, 20 minutes of strength exercises (using ankle weights), and a cool down period. The warm-up and cool down sessions included flexibility and balance exercises. The subject attendances were recorded down and were used to evaluate the adherence level in the study.

At the same time, subjects were encouraged to conduct physical activity trainings outside of class. Investigators also provided them with instructions on taking records of any physical activities (strengthening, stretching and balancing) done out of class.

Screening, Randomization and Blinding Protocols

A total of 154 subjects were screened for the eligibility to participate in this study and 150 subjects qualified (84 at Tufts University, Boston, MA, United States and 66 at Uppsala University and Karolinska Institutet, Stockholm, Sweden). Subjects were subsequently randomized into two arms but four groups (two groups in each arm) in order to increase the level of blindness in this study. During the study, the combination codes of the randomization groups were only known by the manufacturer (NESTLÉ PTC Konolfingen, Switzerland) and will not be revealed to investigators until the final version of the statistical analysis plan (SAP) is approved and the definition of intention to treat (ITT) analyses and per-protocol (PP) analyses are finalized. For the purposes of analyses described

here, randomization is not considered as a factor in adherence; all participants are considered in a unified analysis.

Inclusion and Exclusion Criteria for the VIVE2 Study

Inclusion and exclusion criteria are described at length in Appendix I. Briefly, participants are community-dwelling men and women of at least 70 years of age who are willing to be randomized and participate in the exercise intervention. They are generally free of major comorbidities including recent major cardiovascular events, major recent surgeries, insulin-dependent diabetes mellitus, cancers and terminal illnesses, but are at risk of mobility limitation as indicated by an SPPB score of no greater than 9.

Measures

SF-36 Measures

The SF-36 measures were taken at the baseline, 3 months visit and 6 months visit. This survey consists of 36 questions evaluating 8 subscales of the general wellbeing. In the end, questions can be recombined into physical component summary (PCS) and mental component summary (MCS) as shown in Table 1. All these subscales and summary scores can be used as independent variables to predict poor adherence in this study. In addition, at the Swedish sites, a translated version of the SF-36 questionnaire was used^{40,41}. The Swedish

version of SF-36 was using a different scaling system but could be normalized through a validated process⁴¹.

TABLE I. Components of SF-36 Tests				
Summaries	Subscales		Items	
Physical Component Summary	Physical Functioning (PF)	3. Vigorous Activities	8. Bend, kneel	
		4. Moderate Activities	9. Walk mile	
		5. Lift, carry groceries	10. Walk several blocks	
		6. Climb several flights	11. Walk one block	
		7. Climb one flight	12. Bathe, dress	
		13. Cut down Time	15. Limited in kind	
	Role-Physical (RP)	14. Accomplished less	16. Had difficulty	
	Bodily Pain (BP)	21. Pain-magnitude	22. Pain-interfere	
	General Health (GH)	1. General health rating	35. Health worse	
		33. Sick easier	36. Excellent	
		34. As healthy as anyone		
Mental Component Summary	Vitality (VT)	23. Pep/life	29. Worn out	
		27. Energy	31. Tired	
	Social Functioning (SF)	20. Social-time	32. Social-extent	
	Role-Emotional (RE)	17. Cut down time	19. Not careful	
		18. Accomplished less		
	Mental Health (MH)	24. Nervous	28. Blue/Sad	
		25. Down in dumps	30. Happy	
		26. Peaceful		

CES-D Measures

In addition to the SF-36 test, the CES-D scale was also given at the baseline, 3 month visit and 6 month visit. This scale consists of 20 questions asking subjects to rate the frequencies of certain emotions, such as loneliness

and sadness, over the last week at a scale from 0 to 3 (in most cases, 0 = rarely, 1 = some or little of the Time, 2 = moderately or much of the time, 3 = most or almost all the time). The total score ranges from 0 to 60 and a higher score is associated with greater depressive symptoms¹⁹. The cutoff CES-D score that aid in identifying individuals at risk for clinical depression, with good sensitivity and specificity, was 16⁴².

400 M Walk Measures

Average walking speed during a 400 M walk test is the primary outcome of the VIVE2 study. This test was also given at the baseline, 3 month visit and 6 month visit. For this study, baseline walking speed and other functional measures are used only for descriptive purposes. Subjects were asked to walk at their normal speed for 10 laps around a 40 meter course, or until they could no longer continue. They were not allowed to walk with the help of an assistive device, such as canes and walkers, or the help of other people. If a subject didn't finish the 400 M walk, the average speed was still calculated and used in the analysis based on the distance the participant traveled and the time spent.

Nutritional Intervention Adherence

During the study, subjects were required to drink the nutritional supplement/placebo every day during a 180-day period. During the three physical activity sessions in a week, subjects were asked to take the drink after the session with the supervision of investigators. During the remaining four days

of the week, subjects were asked to take record down whether they had taken the drinks on a provided log sheet. The supervised intakes and the self-administrating intakes were treated equally in the analyses presented here.

Exercise Program Adherence

As discussed above, subjects were required to participate on a 3-day a week physical activity program. Their attendance was recorded down by study nurses and the records were used to evaluate adherence. As with the nutritional records, subjects were instructed to record down the daily completion of the 3-portion exercise (strength, stretching and balance) during the rest of the week on a provided log sheet.

Definition of Risk of Poor Adherence

In the statistical analysis of this study, poor adherence was evaluated either as a continuous variable, or as a discrete variable. First, the overall risk of exercise/product nonadherence for each subject was calculated based the reported daily completion of exercise activities or product consumptions. This variable was a continuous number which evaluated the probability of non-adhering events to occur.

Then, based on the Statistical Analysis Plan (SAP) of the VIVE2 study, the per-protocol analysis would exclude subjects who completed less than 60% of planned exercise activities or 60% of the nutritional supplements. It means that when subjects' risks of non-adherence were smaller than or equal to 40%, they

would be classified as “Good-adherent” but when the risks increased over 40%, they would be classified as “Poor-adherent”. The mathematical equation for this process is described as below.

$$\text{Poor – Product – Adherent: } \frac{\text{Days with No Product Consumption}}{\text{Total Weeks of the Study} \times 7} > 0.4$$

$$\text{Poor – Exercise – Adherent: } \frac{\text{Absent Days of Physical Training Sessions}}{\text{Total Weeks of the Study} \times 3} > 0.4$$

Statistical Analysis Plan

Primary Analysis

The primary analysis of this study was to assess the association between participants’ baseline scores on SF-36 and CES-D tests and adherence to planned exercise and nutritional activities. Poor adherents were defined as subjects who completed less than 60% of the assigned exercises and product consumption, independently. Since the test scores are numerical while the primary outcome is a binary variable, the associations were examined by multivariate logistic regression. A total of four multivariate models were constructed. The first one was a simple model only adjusted by study location and the second one was adjusted by study, location, age, and sex. Model 3 adjusted study location, age, sex, and 400 M walking speed and Model 4 adjusted all the factors in model 3 plus BMI, vitamin D level models, and Mini-

Nutrition Assessment Short Form (MNA-SF) as nutrition were found to be important factors to mobility in elderly population⁴³.

Secondary Analyses

Risk of Poor Adherence (continuous) Analysis

Following the primary analysis, the first secondary analysis was to assess the effects of baseline tests' scores on the overall probability of non-adhering events to occur. In this case, all of the variables, including the outcome, are continuous variables. As a result, the associations were examined using linear regression analysis. Similarly, four multivariate models were built and adjusted by the same factors in the primary analysis.

SF-36 Subscale Score Analyses

We also investigated the association between each SF-36 subscale score (PF, RP, BP, GH, VT, SF, RE, and MH) and the odds of poor adherence by building a multiple logistic regression models considering each of the 8 subscores independently, along with study site.

Risk-prediction Model for Poor Adherence

Finally, all of these factors were considered in risk-prediction models for poor adherence. This consisted of multiple logistic regression models considering the 8 subscales of SF-36 and the CES-D score, with

control for age, sex, study center, BMI, vitamin D level, MNA-SF score, 400 M walking speed, and SPPB score.

RESULTS

Description of trial sample

Baseline characteristics for the 150 participants in the VIVE2 study are shown in Table II; baseline characteristics stratified by randomization groups are listed in Table A.1 in Appendix II. In this study, the subject's age ranged from 69 to 100 years and have a sample mean \pm standard deviation (SD) age of 78 ± 5 years. About 54% of the participants were male. This percentage is higher than the national average (41%) in the United States in 2010⁴⁴.

TABLE II. Baseline Overall Descriptive Characteristics (N=150)			
Demographics			
Age, yr	78 (5)	BMI, kg/m ²	28.2 (3.6)
Men (N, %)	81 (54%)	Vitamin D, nmol/L	42.7 (9.7)
US site (N, %)	84 (56%)		
Baseline Scores		Tests	
MNA-SF Score	13.3 (1.2)	400 M Speed, m/s	0.91 (0.16)
SF-36 Scores			
SF-36 Summary Scores			
PCS	45.7 (8.6)	MCS	51.6 (9.3)
SF-36 Subscale Scores			
Vitality	61.5 (18.5)	Role-physical	64.1 (33.5)
Physical Functioning	67.2 (19.7)	Role-emotional	72.6 (31.0)
Bodily Pain	68.9 (22.6)	Social Functioning	83.3 (20.9)
General Health	67.1 (17.4)	Mental Health	77.8 (16.7)
CES-D Score	11.6 (7.2)		
Abbreviation: BMI, body mass index; MNA-SF, Mini Nutritional Assessment Short-Form; SF-36, Short Form Health Survey; PCS, Physical Component Summary; MCS, Mental Component Summary; CES-D, Center for Epidemiological Study Depression Scale.			

The average baseline BMI level was 28.2 ± 3.6 kg/m², and most of the participants met a crude definition of "overweight" (BMI ≥ 25.0 kg/m²)⁴⁵ at

baseline. The average 25(OH) vitamin D level was 42.7 ± 9.7 nmol/L and 76.2 % met the commonly used criterion for vitamin D insufficiency ($25(\text{OH}) \text{ D} < 50$ nmol/L)⁴⁶. The baseline MNA-SF test indicated that most subjects had normal nutritional status because their test score dropped within the normal range of 12 to 14 points⁴⁷.

In this study, the average PCS score of SF-36 test (45.7 ± 8.6) were below the 1998 national average but the average MCS score (51.6 ± 9.3) were above the average as the national averages for both PCS and MCS were normalized to 50.0 ± 10.0 ⁴⁸. In addition, after normalization with the national average data, 5 out of the 8 subscales, namely physical function, bodily pain, general health, role-physical and role emotional were below national average. The other 3 subscale, including vitality, social functioning and mental health, were above national average. Most subjects did not meet the most commonly used criterion for prevalent depressive symptoms ($\text{CES-D} > 16$).¹⁹

Distribution of Adherence statistics

Center-specific distributions of adherence are depicted in Figure 1. A total of 12% of individuals met the criterion for poor adherence with the exercise program. Adherence to the nutritional intervention was higher, with only 3% of individuals meeting the criterion for nutrition non adherence. The mean \pm SD proportion of visits at which participants were non-adherent for exercise was 0.21 ± 0.14 , as compared to 0.07 ± 0.09 for nutrition.

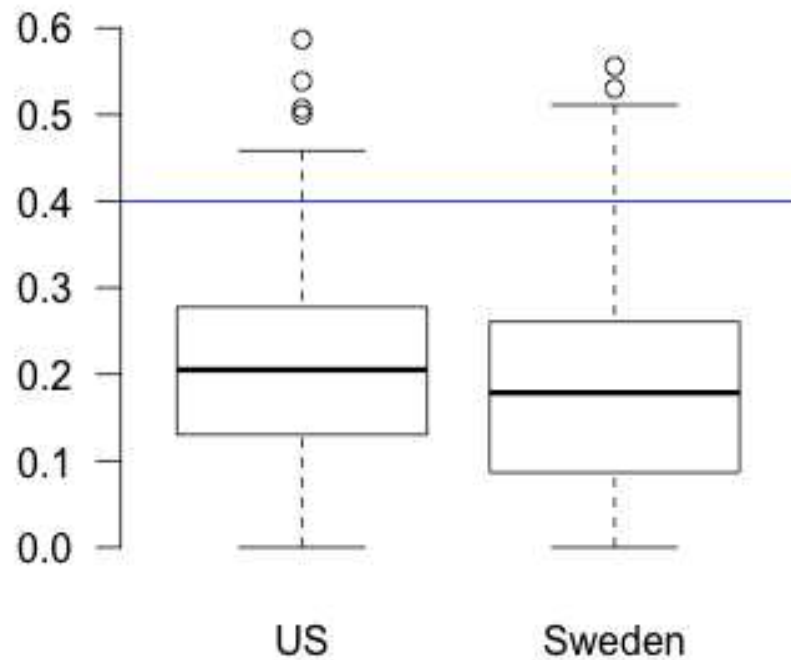


Figure 1. Boxplots showing distribution of exercise non-adherence score by site. According to the protocol, participants with 40% proportionate non-adherence (above blue line) are considered noncompliant.

Association of SF-36 and CES-D Scores with Poor Adherence

The results of the primary analysis are presented in Table III. The general trend is that both SF-36 summary scores were associated with a slightly lower risk of being categorized as exercise or product poor adherents while there seems to be no association between CES-D scores and odds of poor adherence. Based on the fully adjusted Model 4, the odds ratio for the association between

PCS and poor exercise adherence was 0.91. In other words, for every one point positive difference in the PCS score, an individual has a model-estimated -9% multiplicative decrease in the odds of nonadherence (95% CI: -0.19 to 0.01). None of these findings were, however, statistically significant, likely due in part to the limited sample size.

TABLE III. Multivariate Logistic Regression Models Demonstrating the Association of Baseline SF-36 and CES-D Scores with Poor Adherence among Subjects in the VIVE2 Study				
	Model 1 ^a OR (95% CI)	Model 2 ^b OR (95% CI)	Model 3 ^c OR (95% CI)	Model 4 ^d OR (95% CI)
Poor Exercise Adherence				
PCS (SF-36)	0.93 (0.84-1.02)	0.93 (0.84-1.02)	0.92 (0.83-1.02)	0.91 (0.81-1.01)
MCS (SF-36)	0.97 (0.87-1.07)	0.96 (0.87-1.06)	0.96 (0.86-1.06)	0.95 (0.84-1.07)
CES-D	0.99 (0.86-1.12)	0.99 (0.86-1.12)	0.98 (0.85-1.12)	1.00 (0.85-1.19)
Poor Product Adherence				
PCS (SF-36)	0.96 (0.83-1.14)	0.96 (0.82-1.13)	0.95 (0.81-1.13)	0.98 (0.82-1.16)
MCS (SF-36)	0.98 (0.84-1.14)	0.98 (0.84-1.14)	0.98 (0.83-1.15)	0.95 (0.79-1.12)
CES-D	0.99 (0.81-1.20)	1.00 (0.81-1.21)	0.99 (0.81-1.20)	1.00 (0.78-1.22)
How to read the table: one point increase in the test scores would affect the odds of being categorized as Poor Adherence by a certain value. Abbreviations: OR, odds ratio; CI, confidence interval; PCS, physical component summary; MCS, mental component summary; SF-36, Short Form Health Survey; CES-D, Center for Epidemiologic Studies Depression Scale. ^a Model 1: associations between test scores and poor adherence adjusted only by study sites. ^b Model 2: associations adjusted sites, age and sex. ^c Model 3: associations adjusted by sites, age, sex, and walking speed. ^d Model 4: associations adjusted by sites, age, sex, walking speed, BMI, vitamin D level models and MNA-SF score.				

Association of Test Scores with Continuous Measures of Nonadherence

When non-adherence was quantified as a continuous outcome (see Methods), the findings generally agreed with the findings in the primary analysis. Overall, higher PCS and MCS scores were insignificantly associated with lower risks of non-adhering events in both exercise completion and nutrition. In addition,

all of the prediction models showed that CES-D was also insignificantly associated with a lower risk of exercise non-adhering events but seemed to have no association with the risk of product consumption non-adhering events.

TABLE IV. Multivariate Linear Regression Models Demonstrating the Association of Baseline SF-36 and CES-D Scores with Risk of Nonadhering Events (%) among Subjects in the VIVE2 Study				
	Model 1 ^a RR (95% CI)	Model 2 ^b RR (95% CI)	Model 3 ^c RR (95% CI)	Model 4 ^d RR (95% CI)
Risks of Nonadhering Exercise Events to Occur (%)				
PCS (SF-36)	0.88 (0.59-1.31)	0.89 (0.60-1.33)	0.92 (0.60-1.41)	0.84 (0.54-1.29)
MCS (SF-36)	0.98 (0.64-1.51)	0.97 (0.63-1.50)	0.96 (0.61-1.49)	0.92 (0.57-1.48)
CES-D	0.92 (0.53-1.59)	0.91 (0.52-1.58)	0.86 (0.49-1.51)	0.86 (0.45-1.61)
Risks of Nonadhering Product Consumption Events to Occur (%)				
PCS (SF-36)	0.84 (0.66-1.07)	0.87 (0.68-1.11)	0.85 (0.66-1.09)	0.85 (0.65-1.11)
MCS (SF-36)	0.95 (0.76-1.20)	0.92 (0.73-1.16)	0.91 (0.72-1.15)	0.90 (0.70-1.17)
CES-D	1.06 (0.80-1.41)	1.02 (0.76-1.35)	0.99 (0.74-1.33)	0.98 (0.71-1.36)
How to read the table: one point increase in the test scores would affect the percentage risks of nonadhering events by a certain value. Abbreviations: RR, relative risk; CI, confidence interval; PCS, physical component summary; MCS, mental component summary; SF-36, Short Form Health Survey; CES-D, Center for Epidemiologic Studies Depression Scale. ^a Model 1: associations between test scores and poor adherence adjusted only by study sites. ^b Model 2: associations adjusted sites, age and sex. ^c Model 3: associations adjusted by sites, age, sex, and walking speed. ^d Model 4: associations adjusted by sites, age, sex, walking speed, BMI, vitamin D level models and MNA-SF score.				

Associations of SF-36 Subscales with Poor Adherence

The SF-36 subscale analyses also reported no association or very weak association between each SF-36 subscale and the odds of poor adherence. In most cases, the odds ratios crossed 1 and the associated p values were large, which indicated relatively little evidence against the null hypothesis of no association. The only one significant association was found between physical

functioning and poor exercise adherence. Even in this case, the association was relatively weak and the odds ratio was found to be 0.95. In other words, for every one point increase in the physical functioning scores (ranging from 0 to 100), an individual would have an estimated 5% lesser risk to being categorized as poor adherent.

Table V. Associations between each SF-36 Subscale and Poor Adherence				
SF-36 Subscale	Poor Exercise Adherence		Poor Product Adherence	
	OR (95%CI)	P value	OR (95%CI)	P value
Physical Functioning (PF)	0.95 (0.92-0.98)	0.01	1.00 (0.95-1.08)	0.88
Role-physical (RP)	0.99 (0.97-1.01)	0.28	1.01 (0.97-1.05)	0.80
Bodily Pain (BP)	1.00 (0.97-1.02)	0.56	0.97 (0.91-1.02)	0.24
General Health (GH)	0.99 (0.96-1.03)	0.74	1.01 (0.95-1.09)	0.76
Vitality (VT)	0.98 (0.94-1.01)	0.23	0.97 (0.91-1.04)	0.40
Social Functioning (SF)	0.99 (0.96-1.02)	0.43	0.99 (0.94-1.05)	0.63
Role-emotional (RE)	1.00 (0.98-1.02)	0.71	1.01 (0.97-1.07)	0.70
Mental Health (MH)	0.99 (0.96-1.03)	0.53	1.00 (0.94-1.08)	0.93
How to read the table: one point increase in the test scores would affect the odds of being categorized as Poor Adherence by a certain value. Abbreviations: OR, odds ratio; CI, confidence interval; SF-36, Short Form Health Survey.				

Risk Prediction Model for Poor Adherence

Table VI below only demonstrates the factors affecting the risk of poor exercise adherence. (As noted in Figure 1, the product adherence rate in the VIVE2 study was very high. As a result, the fully adjusted model could not be built for risk of poor product adherence because the degrees of freedom were insufficient.)

Table VI. Multivariate Logistic Regression Model Assessing the Associations of Each SF-36 Subscale and Poor Adherence^a

SF-36 Subscales	Poor Exercise Adherence	
	OR(95%CI)	P value
Physical Functioning (PF)	0.93 (0.88-0.98)	0.01
Role-physical (RP)	0.99 (0.96-1.01)	0.29
Bodily Pain (BP)	1.00 (0.96-1.03)	0.93
General Health (GH)	0.97 (0.91-1.03)	0.30
Vitality (VT)	0.97 (0.92-1.02)	0.23
Social Functioning (SF)	0.96 (0.91-1.01)	0.09
Role-emotional (RE)	1.00 (0.98-1.03)	0.83
Mental Health (MH)	0.92 (0.85-0.99)	0.03
CES-D Score	1.04 (0.94-1.16)	0.45

How to read the table: every one point increase in the test scores (for walking speed, every 0.1 m/s increase) would affect the odds of being categorized as Poor Adherence by a certain value. Abbreviations: OR, odds ratio; CI, confidence interval; SF-36, Short Form Health Survey. ^aThis is a multivariate model adjusted by study sites, age, sex, 400 M walking speed, BMI, vitamin D level and MNA-SF score.

In this model, physical functioning and mental health were found to be significantly associated with a decreased risk of being categorized as poor exercise adherents. Every one point cross-sectional increase in PF and MH would decrease the odds of poor exercise adherence by 7% and 8%. The other 6 subscales of the SF-36 test and the CES-D test are found to have no association with poor adherence.

Other Factors

In a model that considered physical functioning, mental health, age, sex (female), study centers, BMI and MNA-SF, subjects completed the study in the United States were reported to be nearly 3 times (OR: 2.97, 95% CI: 0.64-15.69) likely to become poor adherents while females reduced the risk by 61% (OR:

0.39, 95% CI: 0.07-1.72). Both of these findings, however, were statistically nonsignificant. Other factors, including age, BMI and MNA-SF score were found to be irrelevant. The detailed values are shown in Table A2 in Appendix II.

DISCUSSION

Interpretation of the Results

This study examined the possible associations of baseline SF-36 and CES-D scores with the in-study risk of poor adherence. Generally, the associations were found to be relatively weak. In the primary analysis, all the four logistic regression models failed to provide statistically significant association but indicated a general trend that both PCS and MCS scores were associated a slightly lower risk of being categorized as poor adherents while CES-D score seemed to have no association.

These findings did not agree with the study hypothesis, which was made based on literature review, intuition, and the results of a previous study²⁰. That study assessed the association between baseline SF-36 scores and risk of poor adherence among subjects with poorly controlled hypertension. It was reported PCS was significantly associated with an increased risk of poor adherence while MCS had no association with poor adherence. One of the possible explanations was that people with poorly controlled hypertension may have comparatively limited concerns with their mobility ability. As a result, physical functioning limitations might be less of a barrier for them to come to study sites and thus limiting their risk of becoming nonadhering due to physical limitations. By contrast, the subjects in the VIVE2 study were chosen specifically for risk of mobility disability. Occasional mobility issues, therefore, could become a reason why they could not present the assigned physical training sessions and increase the risk of

poor adherence. In this specific population, however, it would be reasonable that a high mobility ability, which was indicated by a higher PCS score, could predict a lower risk of poor adherence. Another possible explanation to the difference between the study finding and study hypothesis was that the hypertension study was conducted in Baltimore, MD, while the VIVE2 study was conducted in Boston, MA and Sweden. The location could play an important role here and affect the associations as many social-economic factors and health care policies, such as medical insurance policies, were different.

Our secondary analyses also found that both physical functioning and mental health were significantly associated with a decreased risk of poor adherence. In fact, in the hypertension study, these two subscale scores were also the only two that had significant associations with risk of poor adherence. The difference was that the hypertension study reported that physical functioning score was positively associated with the risk. The hypothesized explanations to this difference should be the same as explained in the last paragraph. In addition, it should be noted that both studies reported that the mental health subscale in the SF-36 test was associated with a lower risk of poor adherence. In other word, when subjects were mentally healthy, they would have a larger chance to follow the study protocol.

The negative association of generalized mental health and risk of poor adherence, however, could not be confirmed by the CES-D study results in this

study. In all of the regression models, the association between CES-D and risk of poor adherence were not statistically significant but trend to be negative. This would imply that increased depressive symptoms would be associated with better compliance, which seems unlikely, and previous literature reported as well that CES-D was associated with a higher risk of poor adherence²¹. We conclude, therefore, that the weak trend toward negative associations between CES-D and poor adherence found in this study were probably chance events and should be disregarded.

In a model considering physical functioning, mental health, age, sex (female), study centers, BMI and MNA-SF, it was also found that study center and sex have large but insignificant associations with poor adherence. Regarding study sites, it was found that subjects who completed study at Tufts University in the United States had a significant larger risk of poor adherence than subjects in Sweden, which agreed with exploratory results (Figure 1). The difference between the two study sites could be caused by many possible factors. It could be possible that people in the United States had a larger risk of being poor adherent. It could also be possible that the study investigators in Sweden pursued more aggressively the participants' adherence. At this time, explanation for the (nonsignificant) difference is not clear. Because the design was stratified by study center, however, these modest differences reinforce the fact that

presentation of the data (and all models) should control for site, which we have done here.

There also was some nonsignificant suggestion that female sex may be associated with a lower risk of poor adherence. Reasons for this are unknown and even the current scientific evidences on this association are usually contradictory and case-specific¹. Since the association found here was also insignificant, it was probably true that it was caused by random error or some indefinable factors.

Study Limitations

Some limitations of this analysis should be recognized. The first and most important limitation is the analyses were performed based on preliminary results from the VIVE2 study with 112 participants records (out of 150 planned). Thus results must be regarded as preliminary. The lack of full adherence data raises the concern that analyses may be underpowered. However, since most of the associations found in this study were null or weak associations, the impact on the qualitative conclusions of this study are probably small. In a case where effects were more substantial or suggested a compelling overall pattern, the potential for type-II error would be more compelling.

Secondly, the VIVE2 study itself was designed and organized in such a way that the risk of non-adherence had already been largely reduced. Subjects were asked to participate in a 3-time per week exercise training session so study

investigators have enough time to communicate and build strong binds with them. In fact, the general adherence levels in the VIVE2 study were very high. Subjects consumed 92.1% of the assigned nutrition supplements and attended 77.8% of the required exercise sessions. It was definitely great for the VIVE2 project but for this analysis, it brought up a limitation that the adherence data were not collected in the “natural” status. Many possible non-adhering cases might have been eliminated before they could occur.

This raises a third major limitation which is the fact that reason for non-adherence is not known. Given the design, it seems plausibly that most of the non-adhering events that really occurred were probably those “unavoidable” random events, such as family businesses and doctor’s appointments. As a result, the risk of poor adherence might trend to be random as well and potential associations eliminated by the design, but this cannot be known for sure.

A less concerning fact was that, as stated in the result section, the sex distribution in this study was different from the one in the general population. In this study, most of the participants were male while the majority of the general population aged over 70 years were actually female⁴⁴. The reason why this difference existed remained unclear. A hypothesized explanation was that it was more attractive for elderly men than elderly women to participate in a study involving long time physical activity trainings. Past studies had shown that on average, men have a relative higher level of physical activities than women in

elderly population aged over 60 years⁴⁹. If this theory was true, the sex distribution of this study's enrollment reflected the general acceptances of this nutrition supplements plus physical exercise therapy between sexes in the real world. In addition, no matter what caused the difference, the effect of sex was adjusted in 3 out of 4 regression models in this study. As a result, it probably had a limited impact on the study conclusion.

Finally, the lack of depressive symptoms among subjects in the VIVE2 study limited the ability to detect the possible association between level of depression and risk of poor adherence. Based on their MCS score, most of the people who chose to participant in this study were mentally healthy. The average MCS score was even higher than the national average. It could be due to the recruiting methods of the VIVE2 study because subjects were not randomly drawn from a community.

In conclusion, this study found that in a well-organized clinical study among elderly people with mobility limitations, baseline SF-36 summary scores (PCS and MCS) were nonsignificantly associated with a slightly lower risk of poor adherence, while baseline CES-D score was found to have no association with the risk of adherence. Out of the 8 SF-36 subscales, only physical functioning was found to be significantly associated with a lower risk of poor exercise adherence. None of the subscale could predict the risk of poor product adherence. But these conclusions are limited by fact that the study sample is

relatively healthy and the trial design mitigates the risks that our analysis examined.

Even though this study reported null or very weak associations of SF-36 and CES-D scores and risk of poor adherence, it might still be true that the risk had already been largely reduced in this explanatory study. In the future, similar analyses could be done in some pragmatic clinical studies, which assess the effectiveness of an intervention under usual (instead of ideal) condition. In those pragmatic trials where the level of interaction between investigators and subjects was low, it might be easier to reveal the association of quality of life or level of depression and risk of poor adherence. If such an association exists, it might be a good idea to define different adherence management strategies for different situations so investigators could pick the most cost-efficient one based on their budgets and need.

APPENDIX I. Extra Terminologies

Inclusion Criteria for the VIVE2 Study

- Male and Female aged 70 years or older
- Community dwelling
- SPPB Score ≤ 9
- Willingness to be randomized and com to the laboratory for 6 months
- Body Mass Index (BMI) $\leq 35 \text{ kg/m}^2$
- Mini-mental state examination (MMSE) ≥ 24
- Serum 25 (OH) D between 22.5 – 60 nmol/L
- Having obtained his/her informed consent
- Able to complete 400 M walk within 15 min

Exclusion Criteria for the VIVE2 Study

- Acute or terminal illness
- Current regular use (more than 1 time per week) of high protein oral nutritional supplements (e.g. Nestlé Boost®, Exceed® etc...)
- Current use of Vitamin D supplements (more than 800 IU per day)
- Myocardial infarction in previous 6 months, symptomatic coronary artery disease, or congestive heart failure
- Upper or lower extremity fracture in previous 6 months
- Concentration of Hemoglobin $< 10 \text{ g/dL}$

- Estimated glomerular filtration rate (GFR) < 30 mL/min (severe decrease in GFR or kidney failure)
- Uncontrolled hypertension (> 150/90 mm Hg)
- Neuromuscular diseases and drugs which affect neuromuscular function
- Hormone replacement therapy
- Insulin-dependent diabetes mellitus
- Uncontrolled non-insulin-depend diabetes (fasting glucose > 200 mg/dL)
- Milk protein allergy
- Major surgery in the past 6 months (requiring general anesthesia)
- Other significant co-morbid disease that would impair ability to participate in the exercise-based intervention, e.g. renal failure on hemodialysis, severe psychiatric disorder (e.g. bipolar, schizophrenia)
- Excessive alcohol usage (> 14 drinks per week)
- Participation in moderate intensity physical activity > 20 minutes/week
- Inability to communicate due to severe, uncorrected hearing loss or speech disorder
- Severe visual impairment (if it precludes completion of assessments and/or intervention)
- Wheelchair bound
- Severe progressive, degenerative neurologic disease

- Severe rheumatologic or orthopedic diseases, e.g., awaiting joint replacement, active inflammatory disease
- Terminal illness with life expectancy less than 12 months, as determined by a physician
- Cancer requiring treatment in the past three years, except for non-melanoma skin cancers or cancers that have clearly been cured or in the opinion of the investigator carry an excellent prognosis
- Severe pulmonary disease, requiring either steroid pills, injections or the use of supplemental oxygen
- Severe cardiac disease, including New York Heart Association (NYHA) Class III or IV congestive heart failure, clinically significant aortic stenosis, history of cardiac arrest, use of a cardiac defibrillator, or uncontrolled angina
- Patient who cannot be expected to comply with treatment, as decided by the Principal Investigator and study physician
- Conditions not specifically mentioned above may serve as criteria for exclusion at the discretion of the clinical site Principal Investigator and/or study physician

APPENDIX II. Extra Tables

TABLE A1. Baseline Descriptive Characteristics Among Randomization Groups				
Variables (mean, SD)	Group A (n_A=39)	Group B (n_B=37)	Group C (n_C=37)	Group D (n_D=37)
Demographics				
Age, yr	76.6 (5.3)	77.1 (4.4)	76.8 (5.4)	79.4 (6.0)
Female (N, %)	18 (46%)	17 (46%)	17 (46%)	17 (46%)
Americans (N, %)	20 (51%)	20 (54%)	21 (57%)	23 (62%)
BMI, kg/m ²	28.3 (4.1)	28.6 (3.8)	27.3 (3.1)	28.3 (3.5)
Vitamin D, nmol/L	41.5 (9.4)	42.8 (11.0)	44.4 (9.8)	42.0 (8.9)
Baseline Tests Scores				
MNA-SF Score	13.2 (1.2)	13.6 (0.7)	13.4 (1.2)	13.1 (1.7)
400 M Speed, m/s	0.91 (0.17)	0.93 (0.14)	0.93 (0.17)	0.86 (0.16)
SF-36 Scores				
SF-36 Summary Scores				
PCS	45.7 (8.6)	44.8 (7.8)	45.8 (7.6)	45.4 (7.7)
MCS	52.0 (9.4)	53.7 (8.4)	50.9 (9.7)	49.9 (9.5)
SF-36 Subscale Scores				
Vitality	62.0(19.9)	63.6(15.3)	61.0(19.4)	59.1(19.6)
Physical Functioning	65.2(22.5)	67.6(19.0)	69.4(16.2)	66.7(20.8)
Bodily Pain	72.6(21.9)	66.9(23.4)	66.7(22.4)	69.3(23.4)
General Health	65.0(16.8)	66.2(18.9)	69.8(17.2)	67.5(17.2)
Role-physical	67.3(32.8)	63.0(34.6)	65.1(33.1)	60.7(34.8)
Role-emotional	80.0(29.2)	73.1(30.1)	73.4(29.5)	63.3(34.1)
Social Functioning	82.9(24.2)	84.6(19.9)	83.3(21.1)	82.1(18.3)
Mental Health	75.9(18.6)	80.2(15.4)	77.3(15.7)	77.8(17.2)
CES-D Score	12.2 (7.3)	9.6 (5.6)	12.1 (8.2)	12.5 (7.2)

Table A2. Multivariate Logistic Regression Model Assessing Other Factors Affecting Risk of Poor Adherence		
	Poor Exercise Adherence	
	OR(95%CI)	P value
Physical Functioning	0.93 (0.88-0.97)	<<0.01
Mental Health	0.98 (0.93-1.02)	0.26
Age	0.97 (0.84-1.10)	0.65
Female	0.39 (0.07-1.72)	0.23
US Center	2.97 (0.64-15.69)	0.17
BMI	0.95 (0.74-1.19)	0.65
MNA-SF	0.96 (0.60-1.74)	0.88
How to read the table: every one point increase in the test scores would affect the odds of being categorized as Poor Adherence by a certain value. Abbreviations: OR, odds ratio; CI, confidence interval; SF-36, Short Form Health Survey.		

BIBLIOGRAPHY

1. Sabaté E. *Adherence to Long-Term Therapies Evidence for Action*. World Health Organization; 2003.
2. Osterberg L, Blaschke T. Adherence to medication. *The New England journal of medicine* 2005;353(5):487-97.
3. Haynes RB, McDonald H, Garg AX, Montague P. Interventions for helping patients to follow prescriptions for medications. *Cochrane database of systematic reviews (Online)* 2002:CD000011.
4. Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA: the journal of the American Medical Association* 2000;283:2701-11.
5. Ferguson JM, Taylor CB, eds. *The Comprehensive Handbook of Behavioral Medicine*. Dordrecht: Springer Netherlands; 1980.
6. Facilitating Treatment Adherence - A Practitioner's Guidebook.
7. Lin J, Sklar GE, Oh VM Sen, Li SC. Factors affecting therapeutic compliance: A review from the patient's perspective. *Therapeutics and Clinical Risk Management* 2008;4:269-286.
8. Lim TO, Ngah BA. The Mentakab Hypertension Study project. Part II--Why do hypertensives drop out of treatment? *Singapore Medical Journal* 1991;32:249-251.
9. Buck D, Jacoby A, Baker GA, Chadwick DW. Factors influencing compliance with antiepileptic drug regimes. *Seizure* 1997;6:87-93.
10. Cochrane GM. Compliance in asthma. In: *European Respiratory Review*. Vol 8.; 1998:348-350.
11. Kyngas H, Rissanen M. Support as a crucial predictor of good compliance of adolescents with a chronic disease. *Journal of clinical nursing* 2001;10:767-774.
12. Kyngäs H. Compliance of adolescents with chronic disease. *Journal of clinical nursing* 2000;9:549-556.

13. Kyngäs H. Compliance of adolescents with diabetes. *Journal of pediatric nursing* 2000;15:260-267.
14. Carney RM, Freedland KE, Eisen SA, Rich MW, Jaffe AS. Major depression and medication adherence in elderly patients with coronary artery disease. *Health psychology* □: official journal of the Division of Health Psychology, American Psychological Association 1995;14:88-90.
15. Horne R. Patients' beliefs about treatment: The hidden determinant of treatment outcome? Editorial. *Journal of Psychosomatic Research* 1999;47:491-495.
16. Horne R, Weinman J. Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. *Journal of Psychosomatic Research* 1999;47:555-567.
17. Abbott PJ, Moore B, Delaney H, Weller S. Retrospective analyses of additional services for methadone maintenance patients. *Journal of Substance Abuse Treatment* 1999;17:129-137.
18. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical care* 1992;30:473-483.
19. Radloff LS. The CES-D Scale: A Self Report Depression Scale for Research in the General. *Applied Psychological Measurement* 1977;1:385-401.
20. Konerman M, Weeks KR, Shands JR, et al. Short Form (SF-36) Health Survey Measures Are Associated With Decreased Adherence Among Urban African Americans With Severe, Poorly Controlled Hypertension. *Journal of Clinical Hypertension* 2011;13:385-390.
21. Farley J, Miller E, Zamani A, et al. Screening for hazardous alcohol use and depressive symptomatology among HIV-infected patients in Nigeria: prevalence, predictors, and association with adherence. *Journal of the International Association of Physicians in AIDS Care (Chicago, Ill. □: 2002)* 9(4):218-26.
22. SF Health Surveys - Optum.com.
23. Gotlib IH, Hammen CL. *Handbook of Depression, Second Edition*. Guilford Press; 2008:708.

24. Kazis LE, Friedman RH. Improving medication compliance in the elderly. Strategies for the health care provider. *Journal of the American Geriatrics Society* 1988;36(12):1161-2.
25. Kelly-Hayes M, Jette AM, Wolf PA, D'Agostino RB, Odell PM. Functional limitations and disability among elders in the Framingham Study. *American Journal of Public Health* 1992;82(6):841-845.
26. Melzer D, Gardener E, Guralnik JM. Mobility disability in the middle-aged: cross-sectional associations in the English Longitudinal Study of Ageing. *Age and ageing* 2005;34(6):594-602.
27. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. *JAMA* 2011;305(1):50-8.
28. Gardener EA, Huppert FA, Guralnik JM, Melzer D. Middle-aged and mobility-limited: prevalence of disability and symptom attributions in a national survey. *Journal of general internal medicine* 2006;21(10):1091-6.
29. Basaria S, Coviello AD, Travison TG, et al. Adverse events associated with testosterone administration. *The New England journal of medicine* 2010;363(2):109-22.
30. O'Hartaigh B, Pahor M, Buford TW, Dodson JA, Forman DE, Gill TM. Physical activity and resting pulse rate in older adults: Findings from a randomized controlled trial. *American heart journal* 2014;168(4):597-604.
31. Merom D, Cumming R, Mathieu E, et al. Can social dancing prevent falls in older adults? a protocol of the Dance, Aging, Cognition, Economics (DAnCE) fall prevention randomised controlled trial. *BMC public health* 2013;13:477.
32. Ni M, Mooney K, Richards L, et al. Comparative impacts of Tai Chi, balance training, and a specially-designed yoga program on balance in older fallers. *Archives of physical medicine and rehabilitation* 2014;95(9):1620-1628.e30.
33. Rejeski WJ, Fielding RA, Blair SN, et al. The lifestyle interventions and independence for elders (LIFE) pilot study: Design and methods. *Contemporary Clinical Trials* 2005;26:141-154.

34. Rejeski WJ, Miller ME, King AC, et al. Predictors of adherence to physical activity in the Lifestyle Interventions and Independence for Elders pilot study (LIFE-P). *Clinical interventions in aging* 2007;2(3):485-94.
35. Vestergaard S, Kronborg C, Puggaard L. Home-based video exercise intervention for community-dwelling frail older women: a randomized controlled trial. *Aging clinical and experimental research* 2008;20(5):479-86.
36. Mead GE, Greig CA, Cunningham I, et al. Stroke: a randomized trial of exercise or relaxation. *Journal of the American Geriatrics Society* 2007;55(6):892-9.
37. Thwaites JH. Practical aspects of drug treatment in elderly patients with mobility problems. *Drugs & aging* 1999;14(2):105-14.
38. Fielding RA, Katula J, Miller ME, et al. Activity adherence and physical function in older adults with functional limitations. *Medicine and science in sports and exercise* 2007;39(11):1997-2004.
39. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *Journal of gerontology* 1994;49(2):M85-94.
40. Sullivan M, Karlsson J, Ware JE. The Swedish SF-36 Health Survey--I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. *Social science & medicine (1982)* 1995;41(10):1349-58.
41. Taft C, Karlsson J, Sullivan M. Performance of the Swedish SF-36 version 2.0. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation* 2004;13(1):251-6.
42. Lewinsohn PM, Seeley JR, Roberts RE, Allen NB. Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychology and aging* 1997;12(2):277-87.
43. Milaneschi Y, Tanaka T, Ferrucci L. Nutritional determinants of mobility. *Current opinion in clinical nutrition and metabolic care* 2010;13(6):625-9.

44. Howden L. *Age and Sex Composition*: 2010. Washington D.C.: U.S. Department of Commerce Economics and Statistics Administration U.S. Census Bureau; 2011.
45. NHLBI Obesity Education Initiative Expert Panel on the Identification E and T of O in A (US). Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults.
46. Thacher TD, Clarke BL. Vitamin D insufficiency. *Mayo Clinic proceedings. Mayo Clinic* 2011;86(1):50-60.
47. Rubenstein LZ, Harker JO, Salva A, Guigoz Y, Vellas B. Screening for Undernutrition in Geriatric Practice: Developing the Short-Form Mini-Nutritional Assessment (MNA-SF). *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 2001;56(6):M366-M372.
48. Ware JE, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *Journal of Clinical Epidemiology* 1998;51(11):903-912.
49. Milanović Z, Pantelić S, Trajković N, Sporiš G, Kostić R, James N. Age-related decrease in physical activity and functional fitness among elderly men and women. *Clinical interventions in aging* 2013;8:549-56.

CURRICULUM VITAE

HAO ZHU

1989

Home 508-545-2982

Cell 217-898-9641

Email haozhu@hsl.harvard.edu

Address 7 Moccasin Path
Natick, MA, 01760

EDUCATION

M.A. in Clinical Investigation

Boston University (BU) Graduation Date: 1/2015

B.S. in Chemical and Biomolecular Engineering

University of Illinois Urbana Champaign (UIUC)

EXPERIENCES

Programmer Analyst I at Hebrew Senior Life Sep, 2013 to current

I am using R and SAS to analyze clinical data.

Summer Student Internship at Hebrew Senior Life Jun, 2014 to August, 2014

My responsibility included doing data cleaning and preliminary data analysis on the phantom dataset of the VIVE2 Study

Teaching Assistant *Fall, 2013/2014*

I assisted in the teaching of biostatistics at BU and improved the average performance of the class.

Undergraduate Researches at UIUC

Research Associate in a lab studying drug delivery vessels 2011-2012

I developed a research thesis on how to increase the titer and transfection efficiency of drug delivering vessels.

Research Assistant in a lab studying apple tree diseases 2010-2011

I assisted postdocs to develop new types of apples trees that can resist to Fire Bright disease

ACTIVITIES

Volunteer

Mass. Eye and Ear

May 2013 to March 2014

I worked on the nursing floor and the surgical center to help nurses transport patients, do medical clearance, organize hospital supply and clean equipment.

Mass. General Hospital

July 2013 to Dec 2013

I worked in the emergency department to help patients when they need

ACS Student Affinities

2009-2012

I volunteered as a general chemistry tutor in the society from 2009 to 2012 and worked as the Tutoring Director from 2010 to 2012 to pair up tutors and students.

IAESTE

2009-2012

I worked as the External Vice President to build up connections between the society and school department.

"Homemade Soap" Project in Engineering Open House

Spring 2010

I set up a display desk showing how to make soap at home

Volunteer of REACT Program

Spring 2009

I volunteered to teach chemistry classes in a primary school.

TECHNIQUES AND SKILLS

Computer Skills: R, R markdown, LaTeX, SAS, Matlab, Mathematica, Excel, Word, Web and Animation Design

Chemistry/Biology Lab Skills: PCR, ELISA, DNA/RNA separation/purification, IR, NMR, UV-Vis, Mass Spectrometry/GC-MS, GC/LC and HPLC,